JOURNAL OF ALGEBRA 31, 45–66 (1974)

The Equivalence of Bilinear Forms

CARL RIEHM*†

Department of Mathematics, University of Notre Dame, Notre Dame, Indiana 46556

Communicated by J. Tits

Received August 8, 1972

Let B: $V \times V \to F$ be a bilinear form on the finite dimensional vector space V over the field F. No assumptions such as symmetry or skew symmetry are made, although we assume that B is nondegenerate. Another such form C: $W \times W \to F$ is equivalent to B if there is an isomorphism φ: $V \to W$ such that $C(\varphi u, \varphi v) = B(u, v)$ for all u and v in V; such a φ is called an isometry of V and W, or more precisely of B and C.

The solution of the equivalence problem for alternating forms is well known, and solutions have also been obtained for symmetric forms, and more generally for hermitian forms and quadratic forms, over special fields.

It seems to be less well known that in most cases the equivalence problem for general bilinear forms has also been "solved" by J. Williamson [10]. His work was extended by G. E. Wall [9] to the case of sesquilinear forms over a division ring. These solutions consist of associating to B first a linear transformation, called the asymmetry of B, and second a sequence $b_1, \ldots, b_m$ of symmetric, alternating, and hermitian forms; then $B \simeq C$ if and only if their asymmetries are similar and $b_1 \simeq c_1$, $b_2 \simeq c_2, \ldots, b_m \simeq c_m$. Thus once one has solved the equivalence problem for the "classical" types of forms over F (and its finite extensions), the general problem of equivalence is also solved.

Williamson's work was motivated originally by problems in linear systems of differential equations and subsequently by the problem of determining the conjugacy classes, or "normal forms" of elements of the classical groups. The connection between these conjugacy classes and equivalence classes of bilinear forms can be found in [9]. The conjugacy problem has also been worked on more directly by many people (e.g., Springer [7], Zassenhaus [11], Cikunov [3], Milnor [5]). I propose to apply techniques similar to those employed in the conjugacy problem, especially by Milnor, to recast the

* With the support of N.S.F. grant GP-29496X1 at the University of Notre Dame.

† Present address: Department of Mathematics, McMaster University, Hamilton, Ontario, Canada.

equivalence theory of bilinear forms in what seems to me to be a clearer and more useful form. Furthermore, this method also leads to a solution in the case (in characteristic 2) previously unsolved in either the bilinear form or conjugacy class theory. The additional difficulties involved in this case are very similar to those encountered in the equivalence theory of integral quadratic forms in dyadic local fields (O'Meara [6]).

The first appendix contains a version of Witt's theorem for pseudo-quadratic forms, from which Witt's theorem for trace-valued hermitian forms is deduced formally. P. Gabriel contributed a second entitled "Degenerate Bilinear Forms".

1. Reduction of the Problem to Three Cases

If $q \in F[X]$ is a polynomial in one variable, the adjoint polynomial $q^* = X^{\deg q} q(X^{-1})$, i.e.,

$$(a_k X^k + a_{k-1}X^{k-1} + \cdots + a_1 X + a_0)^* = a_0 X^k + a_1 X^{k-1} + \cdots + a_{k-1}X + a_k$$

if $a_k \neq 0$. Clearly

$$q^{**} = q \quad \text{if} \quad q(0) \neq 0, \qquad (q_1 q_2)^* = {q_1}^* {q_2}^*. \tag{1}$$

q is called *self-adjoint* if $q^* = q$, *skew adjoint* if $q^* = -q$.

Proposition 1. *Let q be irreducible in $F[X]$. Then q^* is also irreducible, and if q is self or skew adjoint,*

$$\deg q \text{ odd} \Rightarrow q = a(X \pm 1), \quad a \in F,$$
$$\deg q \text{ even} \Rightarrow q \text{ is self adjoint.}$$

Proof. q^* is irreducible by (1). Suppose $q^* = \epsilon q$ with $\epsilon = \pm 1$. If $k = \deg q$ is odd, then q is a linear combination of terms of the form $X^i + \epsilon X^{k-i}$, so $q(-\epsilon) = 0$; thus $X + \epsilon$ is a factor of the irreducible polynomial q, so $q = a(X + \epsilon)$. Since $q(\epsilon) = \epsilon^k q^*(\epsilon) = \epsilon^{1+k} q(\epsilon)$, $\epsilon = 1$ if $\deg q$ is even.

Let B: $V \times V \to F$ be a bilinear form. Assume B is nondegenerate, i.e., $B(u, v) = 0$ for all v implies $u = 0$. This is equivalent to $B(u, v) = 0$ for all u implying $v = 0$ since both are equivalent to $\det(B(v_i, v_j)) \neq 0$ ($v_1, \ldots, v_n$ a basis of V). The linear maps $v \mapsto B(\ , v)$ and $v \mapsto B(v, \ )$ of V into its dual are both isomorphisms and so there is a unique $\sigma \in GL(V)$ such that $B(\ , \sigma v) = B(v, \ )$ for all v, i.e.,

$$B(u, \sigma v) = B(v, u) \qquad \text{for all} \quad u, v \in V. \tag{2}$$

We call σ the *asymmetry* of B. Thus for example $\sigma = 1$ (identity) iff B is symmetric, $\sigma = -1$ iff B is skew symmetric.

If B is degenerate it does not necessarily have an asymmetry, i.e., a $\sigma \in GL(V)$ satisfying (2), nor would an asymmetry be unique. Suppose that B has an asymmetry. Make V into an $F[X]$-module such that $X \cdot v = \sigma v$. Since $B(u, v) = B(\sigma u, \sigma v)$, we have $B(\sigma u, v) = B(u, \sigma^{-1}v)$ and more generally if $q \in F[X]$ has degree k

$$B(q(\sigma)u, v) = B(u, q(\sigma^{-1})v) = B(u, \sigma^{-k}q^*(\sigma)v). \tag{3}$$

Let us say that subspaces U and W of V are *orthogonal* if $B(U, W) = 0$. This does *not* imply in general that W and U are orthogonal.

If U is a submodule of the $F[X]$-module V then the left orthogonal complement $\{v \in V : B(v, U) = 0\}$ of U and the right orthogonal complement $\{v \in V : B(U, v) = 0\}$ are equal since $B(U, v) = B(v, \sigma U) = B(v, U)$; thus in this case W and U are orthogonal iff U and W are orthogonal. In particular the left and right radicals of B (i.e., the two orthogonal complements of V) are equal. Conversely it is easy to show that if the left and right radicals are equal, then B has an asymmetry, although we shall not use this fact.

PROPOSITION 2. *Suppose V (i.e., B) is non-degenerate and $V = U \oplus W$ with U and W orthogonal. Then U and W are nondegenerate subspaces of V. Further W and U are also orthogonal if and only if one of them is an $F[X]$-submodule; and then both are submodules, and we write $V = U \perp W$.*

Proof. The proof of the nondegeneracy of U and W is straightforward and it has already been shown that W and U are orthogonal if one of them is a submodule. Conversely if W and U are orthogonal then since $B(W, U) = B(U, \sigma W) = B(\sigma^{-1}U, W)$, we must have $\sigma W \subseteq W$ and $\sigma^{-1}U \subseteq U$ so both are submodules.

In general, if $V = V_1 \oplus \cdots \oplus V_r$ with V_i and V_j orthogonal for all distinct i and j, we write $V = V_1 \perp \cdots \perp V_r$ and say that V is the *orthogonal* direct sum of $V_1, \ldots, V_r$. It follows from Proposition 2 that each V_i is then a submodule of V if V is nondegenerate.

For each monic irreducible divisor p of the minimal polynomial of σ, the p-primary component $V_p = \{v \in V : p^r v = 0 \text{ for } r \gg 0\}$ is a submodule, and V is their direct sum.

PROPOSITION 3. *V_p and $V_{p'}$ are orthogonal unless $p' = cp^*$ for some $c \in F$.*

Proof. Suppose $p' \neq cp^*$. Then p^* and p' are relatively prime and so $p^*(\sigma)$ is bijective on $V_{p'}$. Choose r large enough so that $p(\sigma)^r = 0$ on V_p; then by (3)

$$0 = B(p(\sigma)^r V_p, V_{p'}) = B(V_p, \sigma^{-kr}p^*(\sigma)^r V_{p'}) = B(V_p, V_{p'}).$$

COROLLARY. *V is the orthogonal direct sum of the subspaces*

$$V_p \oplus V_{p^*} \ (p^* \neq \pm p) \qquad \text{and} \qquad V_p \ (p^* = \pm p).$$

THEOREM 4. *Let B and C be nondegenerate bilinear forms on V and W resp. If V and W are isometric, the asymmetries σ and τ of B and C are similar. Moreover V and W are isometric if and only if for every monic irreducible polynomial p, the following isometries hold:*

$$V_p \oplus V_{p^*} \simeq W_p \oplus W_{p^*} \qquad \text{when} \qquad p^* \neq \pm p,$$
$$V_p \simeq W_p \qquad \text{when} \qquad p^* = \pm p.$$

Proof. If $\varphi: V \to W$ is an isometry, then

$$C(\varphi u, \varphi\sigma v) = B(u, \sigma v) = B(v, u) = C(\varphi v, \varphi u) = C(\varphi u, \tau\varphi v)$$

and so $\varphi\sigma = \tau\varphi$. It follows that σ and τ are similar, and that φ is an $F[X]$-homomorphism whence $\varphi: V_p \tilde{\to} W_p$ and so the necessity of the last statement is proved. The sufficiency is trivial by the corollary of Proposition 3.

Theorem 4 reduces the equivalence problem for nondegenerate forms to 2 cases:

Case I: $V = V_p \oplus V_{p^*}$, $W = W_p \oplus W_{p^*}$, $p^* \neq \pm p$.

Case II: $V = V_p$, $W = W_p$, $p^* = \pm p$.

For reasons to appear later, II breaks up into 2 cases:

Case IIa: $\deg p > 1$, or $\deg p = 1$ and $\operatorname{char} F \neq 2$.

Case IIb: $\deg p = 1$ and $\operatorname{char} F = 2$.

By Proposition 1, $\deg p = 1$ and $p^* = \pm p$ means that $p = X \pm 1$. Case I is easily disposed of:

THEOREM 5. *In Case* I, *$V \simeq W$ if and only if σ and τ (the asymmetries of B and C) are similar.*

Proof. The necessity was proved in Theorem 4. Suppose then that σ and τ are similar. Then $\sigma|_{V_p}$ and $\tau|_{W_p}$ are also similar. Choose $\psi: V_p \tilde{\to} W_p$ to satisfy $\psi\sigma = \tau\psi$ on V_p.

By Proposition 3 both V_p and V_{p^*} are self orthogonal and so B makes V_{p^*} into the dual space of V_p : $\langle u, v\rangle_{V_p} = B(u, v)$ for $u \in V_p$, $v \in V_{p^*}$. Similarly W_{p^*} can be viewed as the dual of W_p, and we let ${}^t\psi: W_{p^*} \to V_{p^*}$ be the transpose of ψ, $\langle \psi u, w\rangle_{W_p} = \langle u, {}^t\psi w\rangle_{V_p}$, and let $\psi^* = {}^t\psi^{-1}$ be the contragredient.

Define

$$\varphi = \psi \oplus \psi^*: V_p \oplus V_{p^*} \to W_p \oplus W_{p^*}.$$

Since these 4 direct summands are self orthogonal, to show that φ is an isometry it suffices to show that $C(\varphi u, \varphi v) = B(u, v)$ if $u \in V_p$, $v \in V_{p^*}$ and if $u \in V_{p^*}$, $v \in V_p$. In the first case

$$C(\varphi u, \varphi v) = C(\psi u, \psi^* v) = B(u, {}^t\psi\psi^* v) = B(u, v)$$

and in the second

$$C(\psi^* u, \psi v) = C(\tau^{-1}\psi v, \psi^* u) = C(\psi\sigma^{-1}v, \psi^* u) = B(\sigma^{-1}v, u) = B(u, v).$$

An alternative proof consists of choosing bases of V_p and W_p in which the restrictions of σ and τ have the same matrix M, enlarging these bases to bases of V and W by adding on the dual bases of V_{p^*} and W_{p^*}, and then showing that the matrices of B and C in these bases are both equal to

$$\begin{pmatrix} 0 & I \\ M^{-1} & 0 \end{pmatrix}.$$

In Case II a further decomposition of $V = V_p$ can be given, although it is not in general uniquely determined. It is shown in the theory of finitely generated torsion modules over $F[X]$ that $V = V_1 \oplus \cdots \oplus V_r$, where V_s is free over $F[X]/(p^s)$ and p^r is the minimal polynomial of σ.

LEMMA 6. *The above decomposition can be chosen so that the V_s are mutually orthogonal*: $V = V_1 \perp \cdots \perp V_r$.

Proof. The proof is by induction on r. Since $r = 1$ is trivial assume $r > 1$. It suffices to show that V_r is a nondegenerate subspace of V. For then $V = V' \oplus V_r$ where V' and V_r are orthogonal; by Proposition 2 $V = V' \perp V_r$ and V' is a submodule and is nondegenerate. By the theory of modules over $F[X]$, the annihilator of V' is of the form (p^s) where $s < r$, so we may apply the induction hypothesis to V' to conclude the proof.

So suppose $0 \neq v \in V_r$. Choose $t \geqslant 0$ so that $p^t v \neq 0 = p^{t+1}v$, and then choose $w \in V_r$ so that $p^{r-1}w = p^t v$ (that w exists can be seen by expressing $p^t v$ in terms of a basis of V_r over $F[X]/(p^r)$). Then if $u \in V_1 \oplus \cdots \oplus V_{r-1}$, $B(p^t v, u) = B(w, \sigma^{-(r-1)k}p^{*r-1}u) = 0$, and so there is an $x \in V_r$ such that

$$0 \neq B(p^t v, x) = B(v, \sigma^{-tk}p^{*t}x)$$

whence $B(v, V_r) \neq 0$, and the proof is complete.

Assume still that $V = V_p$. Then the subring $F[\sigma]$ of $\mathrm{End}_F V$ is $\simeq F[X]/(p^r)$ and so is a local ring with proper ideals $(p(\sigma)) \supset (p(\sigma)^2) \supset \cdots \supset (p(\sigma)^r) = 0$. Let π be any generator of $(p(\sigma))$, and define for $s \geqslant 0$

$$V(s) = \{v \in V : \pi^s v = 0\} = \ker(v \mapsto \pi^s v)$$

where $\pi^0 = 1$. Then $V(s)$ is a submodule of V independent of the choice of the generator π, and it is easy to see that if $V = V_1 \oplus \cdots \oplus V_r$ is a splitting of the above type then $V(s) = V$ if $s \geqslant r$ and

$$V(s) = V_1 \oplus \cdots \oplus V_s \oplus \pi V_{s+1} \oplus \cdots \oplus \pi^{r-s} V_r \tag{4}$$

if $s < r$.

Since

$$p(\sigma^{-1})^r = \sigma^{-rk} p^*(\sigma)^r = 0,$$

the map $\sigma \mapsto \sigma^{-1}$ induces an involution on $F[\sigma]$ taking π to another generator π' of $(p(\sigma))$; then $B(u, \pi v) = B(\pi' u, v)$ for all $u, v \in V$. Suppose $V = V_1 \perp \cdots \perp V_r$, so each V_s is nondegenerate. If $u, v \in V_s$ and $t \leqslant s$ then $B(u, \pi^t v) = B(\pi'^t u, v)$ so the orthogonal complement (in V_s) of $\pi^t V_s$ is $V(t) \cap V_s = \pi^{s-t} V_s$. Using (4), one obtains in a straightforward manner

$$(\pi^t V(s))^\perp = V(t) + \pi^{s-t} V, \tag{5}$$

and furthermore it is easy to see that

$$\begin{aligned}(\pi^{s-1} V(s))^\perp \cap V(s) &= V_1 \perp \cdots \perp V_{s-1} \perp \pi V_s \perp \pi V_{s+1} \perp \cdots \perp \pi^{r-s} V_r \\ &= V(s-1) + \pi V(s+1)\end{aligned} \tag{6}$$

whence

$$V_s/\pi V_s \simeq V(s)/(V(s-1) + \pi V(s+1)) \tag{7}$$

via the map $v + \pi V_s \mapsto v + (V(s-1) + \pi V(s+1))$.

2. Hermitian Forms

Let U be a vector space of dimension m over the field E and suppose that E has an involution ι. Suppose that $\mu \in E$ or $\mathbf{Z}$ and $\mu\mu^\iota = 1$, with the obvious interpretation if $\mu \in \mathbf{Z}$. A map h: $U \times U \to E$ is a μ-hermitian form if it is linear in the first variable, ι-linear in the second i.e., additive and $h(u, av) = a^\iota h(u, v)$, and if $h(u, v) = \mu h(v, u)^\iota$ for all u and v. The set of μ-hermitian forms on U forms a vector space $\mathbf{H}(\iota, \mu)$ over the fixed field E_0 of ι.

If ι is the identity, $\mu = \pm 1$ and we view μ as being in $\mathbf{Z}$ rather than in E. When $\mu = 1$, h is a symmetric bilinear form. When $\mu = -1$ and $\operatorname{char} F \neq 2$, h is alternating; we adopt the convention that when $\operatorname{char} F = 2$, $\mu = -1$ again means that h is alternating.

If $\epsilon = \pm 1 \in \mathbf{Z}$, $\varphi \in \operatorname{End}_E U$ is called *ϵ-adjoint* (with respect to h) if

$$h(\varphi u, v) = h(u, \epsilon\varphi v)$$

for all u and v. The ϵ-adjoint endomorphisms with respect to h form an E_0-subspace $\mathbf{A}(\epsilon)$ of $\operatorname{End}_E U$.

LEMMA 7. (a) *Suppose $\iota \neq$ identity. Then* $\dim_{E_0} \mathbf{H}(\iota, \mu) = m^2$, *and* $\dim_{E_0} \mathbf{A}(\epsilon) = m^2$ *if h is nondegenerate.*

(b) *Suppose $\iota =$ identity, so $\mu = \pm 1 \in \mathbf{Z}$. Then* $\dim_E \mathbf{H}(\iota, \mu) = \frac{1}{2}m(m + \mu)$. *If h is nondegenerate,* $\dim_E \mathbf{A}(\epsilon) = \frac{1}{2}m(m + \epsilon\mu)$ *when* $\operatorname{char} E \neq 2$, *and* $= \frac{1}{2}m(m + 1)$ *when* $\operatorname{char} E = 2$.

Proof. The choice of a basis of U leads to an E_0-isomorphism $h \mapsto H$ from $\mathbf{H}(\iota, \mu)$ to the space of m by m matrices which satisfy ${}^tH = \mu H^\iota$, and which have 0 diagonal in the alternating case $\iota =$ identity and $\mu = -1$ (this additional requirement is really only necessary when $\operatorname{char} E = 2$). This means that the entries above the diagonal can be chosen arbitrarily and then those below the diagonal are completely determined, while each diagonal entry must satisfy $a = \mu a^\iota$, and must be 0 in the alternating case. The formula for $\dim \mathbf{H}(\iota, \mu)$ in (b) follows at once. Suppose $\iota \neq$ identity. Then the E_0-subspace of E of those $a \in E$ satisfying $a = \mu a^\iota$ can be shown to have dimension 1 (note that $a = 1 + \mu$ is in it). Thus $\dim_{E_0} \mathbf{H}(\iota, \mu) = 2 \cdot \frac{1}{2}(m - 1)m + m = m^2$.

Let Φ be the matrix of $\varphi \in \operatorname{End} V$. Then φ is ϵ-adjoint with respect to $h \in \mathbf{H}(\iota, \mu)$ iff ${}^t\Phi H = \epsilon H \Phi^\iota$, which is equivalent to ${}^t(H^\iota\Phi) = \epsilon\mu^\iota(H^\iota\Phi)^\iota$ since ${}^tH = \mu H^\iota$. Unless $\operatorname{char} E = 2$ and $\iota =$ identity, this says that $H^\iota\Phi \in \mathbf{H}(\iota, \epsilon\mu^\iota)$ and so if h is nondegenerate, i.e., H is invertible, $\dim_{E_0} \mathbf{A}(\epsilon) = \dim_{E_0} \mathbf{H}(\iota, \epsilon\mu^\iota)$ and the formulas for $\dim \mathbf{A}(\epsilon)$ follow in this case. When $\operatorname{char} E = 2$ and $\iota =$ identity, the condition is that $H\Phi$ be symmetric, so $\dim_E \mathbf{A}(\epsilon) = \frac{1}{2}m(m + 1)$ and the proof is finished.

Now suppose that E is a finite extension of a field $F \subseteq E_0$. Let $\operatorname{Tr}\colon E \to F$ be a nonzero F-linear map such that $\operatorname{Tr}(a^\iota) = \operatorname{Tr} a$ for all a in E. We take $\operatorname{Tr} =$ identity if $F = E$. If E/F is separable one can take $\operatorname{Tr} = \operatorname{Tr}_{E/F}$, otherwise Tr can be chosen to be $\operatorname{Tr}_{E/E_0}$ followed by any nonzero F-linear map $E_0 \to F$.

If $h \in \mathbf{H}(\iota, \mu)$ we let $\operatorname{Tr}_* h = \operatorname{Tr} \circ h\colon U \times U \to F$. Then $D = \operatorname{Tr}_* h$ is a bilinear form satisfying

$$D(au, v) = D(u, a^\iota v), \qquad D(u, v) = D(v, \mu u) \tag{8}$$

for all u and v. The first condition says that ι is the adjoint with respect to D on $E \subseteq \mathrm{End}_F U$, and the second says that μ is an asymmetry of D. Let $\mathbf{B}(\iota, \mu)$ be the collection of bilinear forms $D\colon U \times U \to F$ satisfying (8). It is a vector space over E_0 if one defines $(aD)(u, v) = D(au, v)$.

THEOREM 8.[1] *The map* $\mathrm{Tr}_* : \mathbf{H}(\iota, \mu) \to \mathbf{B}(\iota, \mu)$ *defined by the commutative diagram*

$$\begin{array}{ccc} U \times U & \xrightarrow{h} & E \\ & \searrow^{\mathrm{Tr}_* h} & \downarrow \mathrm{Tr} \\ & & F \end{array}$$

is an isomorphism of E_0-vector spaces, and the radical of $\mathrm{Tr}_* h$ *= the radical of h.*

Proof. That it is an E_0-homomorphism is easy to see. To show that it is an isomorphism it suffices to prove that, given $D \in \mathbf{B}(\iota, \mu)$, there is a unique $h \in \mathbf{H}(\iota, \mu)$ with trace D.

We begin by showing that there is a unique map $h\colon U \times U \to E$ satisfying

$$\mathrm{Tr} \circ h = D, \qquad h(au, v) = ah(u, v) \qquad \text{for all} \quad u, v.$$

The dual space E^* of the F-vector space E has dimension $|\, E : F \,|$ over F. But it is also a vector space over E via $(b\lambda)(a) = \lambda(ab)$ and hence must have dimension 1 over E. Thus $E^* = E\mathrm{Tr}$. If u and v are vectors in U, the map $a \mapsto D(au, v)$ is in E^* and so there is a unique element in E, call it $h(u, v)$, such that

$$D(au, v) = \mathrm{Tr}(ah(u, v))$$

for all a in E.

If $b \in E$ then $\mathrm{Tr}\, ah(bu, v) = D(abu, v) = \mathrm{Tr}\, abh(u, v)$ so the uniqueness of $h(bu, v)$ implies that $h(bu, v) = bh(u, v)$ so h has the required properties. If g is another such map $U \times U \to E$, then $\mathrm{Tr}\, ag(u, v) = \mathrm{Tr}\, g(au, v) = D(au, v) = \mathrm{Tr}\, ah(u, v)$ for all a, u and v and so again the uniqueness of $h(u, v)$ implies that $g = h$.

We now show that $h \in \mathbf{H}(\iota, \mu)$. To show that h is additive in both variables, one uses the definition of h, the biadditivity of D and additivity of Tr to show, e.g., that

$$\mathrm{Tr}\, ah(u_1 + u_2, v) = \mathrm{Tr}\, a(h(u_1, v) + h(u_2, v))$$

from which additivity in the first variable follows, again by the uniqueness of h.

[1] This theorem is, to a large extent, a special case of Theorem 7.1 in [4].

Similarly using (8) and $\mathrm{Tr} \circ \iota = \mathrm{Tr}$ one gets $h(v, u) = \mu h(u, v)^\iota$, from which, along with linearity in the first variable, one gets semilinearity in the second variable, and so h is a μ-hermitian form as desired.

Since $b = 0$ iff $\mathrm{Tr}\, ab = 0$ for all a in E, $h(u, v) = 0$ for all u in V iff $\mathrm{Tr}\, h(au, v) = 0$ for all a and u. Thus $\mathrm{rad}\, h = \mathrm{rad}\, \mathrm{Tr}_* h$ as desired.

3. Case IIa

In this section $V = V_p$ with $p^* = \pm p$, and also $\mathrm{char}\, F \neq 2$ when $\deg p = 1$ i.e., $p = X \pm 1$. If $\deg p > 1$ then $\deg p$ is even and p is self adjoint by Proposition 1, and so if we define

$$\pi = \sigma^{-(1/2)\deg p} p(\sigma) \in F[\sigma] \subseteq \mathrm{End}_F V,$$

it is easily seen that π is self adjoint with respect to B,

$$B(\pi u, v) = B(u, \pi v).$$

If $p = X \pm 1$ and

$$\pi = \sigma - \sigma^{-1} \in F[\sigma] \subseteq \mathrm{End}_F V$$

then π is skew adjoint with respect to B,

$$B(\pi u, v) = B(u, -\pi v).$$

Define $\epsilon = \pm 1$ to satisfy $B(\pi u, v) = B(u, \epsilon \pi v)$. The other important property of π is that it generates the maximal ideal $(p(\sigma))$ of $F[\sigma]$. Let $E = F(\bar{\sigma})$ be the residue class field $F[\sigma]/(\pi)$. The minimal polynomial of $\bar{\sigma}$ is p, and since $p^* = \pm p$, $\bar{\sigma}^{-1}$ is also a root of p and so E has an involution ι such that $\bar{\sigma}^\iota = \bar{\sigma}^{-1}$. It is the identity iff $\deg p = 1$, and then $\bar{\sigma} = \pm 1$ and $E = F$.

If $1 \leqslant s \leqslant r$ let $V(s) = \ker(v \mapsto \pi^s v)$ as in Section 1 and define a bilinear form B_s on $V(s) \times V(s)$ by $B_s(u, v) = B(\pi^{s-1} u, v)$. It has $\epsilon^{s-1}\sigma$ (restricted to $V(s)$) as asymmetry and radical $(\pi^{s-1}V(s))^\perp \cap V(s) = V(s-1) + \pi V(s+1)$ by (6). Define $\overline{V}_s = V(s)/(V(s-1) + \pi V(s+1))$; it is an $F[\sigma]$-module annihilated by π and so we may suppose that $E \subseteq \mathrm{End}_F \overline{V}_s$. Furthermore there is a nondegenerate bilinear form

$$\bar{B}_s : \overline{V}_s \times \overline{V}_s \to F, \qquad \bar{B}_s(\bar{u}, \bar{v}) = B(\pi^{s-1} u, v)$$

with asymmetry $\epsilon^{s-1}\bar{\sigma} \in E$ such that ι is the adjoint of $\bar{B}_s$ on E : $\bar{B}_s(a\bar{u}, \bar{v}) = \bar{B}_s(\bar{u}, a^\iota \bar{v})$ for $a \in E$. With Tr: $E \to F$ as in the preamble of Theorem 8,

there is by this theorem a unique nondegenerate $\epsilon^{s-1}\bar{\sigma}$-hermitian form

$$h_s : \overline{V}_s \times \overline{V}_s \to E$$

satisfying $\mathrm{Tr} \circ h_s = \bar{B}_s$.

Suppose C: $W \times W \to F$ is another bilinear form with $W = W_p$ and asymmetry similar to σ. We denote the asymmetry of W by σ also and identify the subrings $F[\sigma]$ which it generates in $\mathrm{End}_F V$ and $\mathrm{End}_F W$. Let $g_1, \ldots, g_r$ be the hermitian forms for C defined as were $h_1, \ldots, h_r$ for B.

THEOREM 9 (Case IIa). *The bilinear forms B and C are equivalent over F, $B \underset{F}{\simeq} C$, if and only if $h_1 \underset{E}{\simeq} g_1, \ldots, h_r \underset{E}{\simeq} g_r$.*

Proof. (1) If φ: $V \to W$ is an isometry, then φ is an $F[\sigma]$-isomorphism (see the proof of Theorem 4) and hence induces E-isomorphisms

$$\varphi_s : \overline{V}_s \to \overline{W}_s, \quad \varphi_s(\bar{u}) = \overline{\varphi(u)}.$$

Thus

$$\mathrm{Tr}\, g_s(\varphi_s \bar{u}, \varphi_s \bar{v}) = C(\pi^{s-1}\varphi u, \varphi v) = B(\pi^{s-1}u, v) = \mathrm{Tr}\, h_s(\bar{u}, \bar{v})$$

and so since the map Tr_* in Theorem 8 is injective, $g_s(\varphi_s \bar{u}, \varphi_s \bar{v}) = h_s(\bar{u}, \bar{v})$ so $g_s \simeq h_s$.

(2) To prove the converse we may assume that $V = V_s$. Indeed we write $\overline{V} = V_1 \perp \cdots \perp V_r$ and $W = W_1 \perp \cdots \perp W_r$ by Lemma 6. Now $V_s \simeq V_s/\pi V_s$ by (7) and it is easy to see that this isomorphism is actually an isometry with respect to h_s and the only nonzero hermitian form attached to $B\,|_{V_s \times V_s}$; this latter form is therefore equivalent to the analogous form attached to $C\,|_{W_s \times W_s}$ and so by assumption there is an isometry $V_s \simeq W_s$ (with respect to the respective restrictions of B and C) for each s, and since V and W are the orthogonal direct sums of these spaces, $B \simeq C$ as desired.

(3) To simplify the notation, we drop the subscript $_s$ from V_s, $\overline{V}_s$, h_s, etc. The proof proceeds by showing, by induction on $t = s - 1, s - 2, \ldots, 0$ the existence of an $F[\sigma]$-isomorphism φ: $V \to W$ such that

$$C(\pi^t \varphi u, \varphi v) = B(\pi^t u, v) \tag{9}$$

for all u, v in V. The map φ for $t = 0$ is then the required isometry.

When $t = s - 1$ we let $\bar{\varphi}$: $\overline{V} \to \overline{W}$ be an isometry, $g(\bar{\varphi}\bar{u}, \bar{\varphi}\bar{v}) = h(\bar{u}, \bar{v})$. Take traces to get $\bar{C}(\bar{\varphi}\bar{u}, \bar{\varphi}\bar{v}) = \bar{B}(\bar{u}, \bar{v})$. "Lift" $\bar{\varphi}$ to an $F[\sigma]$-homomorphism φ: $V \to W$, so that $\overline{\varphi u} = \bar{\varphi}\bar{u}$. (This can be done by choosing a basis $v_1, \ldots, v_n$

of V over $F[\sigma]$, defining $\varphi v_1, \ldots, \varphi v_n$ in W so their images in $\overline{W}$ are $\bar{\varphi}\bar{v}_1, \ldots, \bar{\varphi}\bar{v}_n$, and then extending by linearity). By the definition of $\bar{C}$ and $\bar{B}$ we get

$$C(\pi^{s-1}\varphi u, \varphi v) = B(\pi^{s-1}u, v)$$

which is (9) for $t = s - 1$. Furthermore the image in E of $\det \varphi$ (using fixed bases of V and W over $F[\sigma]$) is $\det \bar{\varphi} \neq 0$ and so $\det \varphi$ is a unit and φ is an isomorphism.

We therefore suppose that $t < s - 1$ and that we have found an $F[\sigma]$-isomorphism φ such that

$$C(\pi^{t+1}\varphi u, \varphi v) = B(\pi^{t+1}u, v). \tag{10}$$

(4) We next construct on E_0-homomorphism D: $\mathrm{Hom}_E(\overline{V}, \overline{W}) \to \mathbf{B} = \mathbf{B}(\iota, \epsilon^t\bar{\sigma}) =$ bilinear forms on $\overline{V} \times \overline{V}$ satisfying (8) with $\mu = \epsilon^t\bar{\sigma}$, and show it is surjective. If $\bar{\psi} \in \mathrm{Hom}_E(\overline{V}, \overline{W})$ we may, as before, lift it to $\psi \in \mathrm{Hom}_{F[\sigma]}(V, W)$ so that $\overline{\psi(v)} = \bar{\psi}(\bar{v})$ for all $v \in V$. Define

$$D_{\bar{\psi}} : \overline{V} \times \overline{V} \to F, \qquad D_{\bar{\psi}}(\bar{u}, \bar{v}) = C(\pi^{s-1}\psi u, \varphi v) + \epsilon^{s-t-1}C(\pi^{s-1}\varphi u, \psi v).$$

Since $\pi^s = 0$ and φ and ψ are $F[\sigma]$-homomorphisms, $D_{\bar{\psi}}$ is a well defined bilinear form; and if $a = q(\bar{\sigma}) \in E$ (where $q \in F[X]$), it follows readily from $C(q(\sigma)w_1, w_2) = C(w_1, q(\sigma^{-1})w_2)$ that $D_{\bar{\psi}}(a\bar{u}, \bar{v}) = D_{\bar{\psi}}(\bar{u}, a^\iota\bar{v})$ since $a^\iota = q(\bar{\sigma}^{-1})$. One can also check that $\epsilon^t\bar{\sigma}$ is an asymmetry of $D_{\bar{\psi}}$, so $D_{\bar{\psi}}$ is in $\mathbf{B}$. Furthermore if ψ' is another lifting of $\bar{\psi}$ to V then $\psi(v) \equiv \psi'(v) \bmod \pi V$ and it follows easily that $\bar{\psi} \mapsto D_{\bar{\psi}}$ is a well defined map D: $\mathrm{Hom}_E(\overline{V}, \overline{W}) \to \mathbf{B}$. It is clearly additive and a proof similar to those above shows that it is E_0-linear.

We now determine the kernel of D. $D_{\bar{\psi}} = 0$ iff

$$C(\pi^{s-1}\psi u, \varphi v) = -\epsilon^{s-t-1}C(\pi^{s-1}\varphi u, \psi v) \qquad \text{for all} \quad u, v \in V.$$

Set $\zeta = \psi\varphi^{-1} \in \mathrm{End}_{F[\sigma]}W$. Since φ is onto, this condition is equivalent to

$$C(\pi^{s-1}\zeta u, v) = -\epsilon^{s-t-1}C(\pi^{s-1}u, \zeta v) \qquad \text{for all} \quad u, v \in W.$$

By the definition of $\bar{C}$ this equality is the same as $\bar{C}(\overline{\zeta u}, \bar{v}) = -\epsilon^{s-t-1}\bar{C}(\bar{u}, \overline{\zeta v})$ and this in turn is equivalent to

$$g(\overline{\zeta u}, \bar{v}) = -\epsilon^{s-t-1}g(\bar{u}, \overline{\zeta v})$$

for all u, v in W; for if this last equation did not hold for some particular u and v, one could multiply each side by a suitable $a \in E$ and take traces to get $\bar{C}(\overline{\zeta a u}, \bar{v}) \neq -\epsilon^{s-t-1}\bar{C}(a\bar{u}, \overline{\zeta v})$. We therefore see that $D_{\bar{\psi}} = 0$ iff $\bar{\zeta} = \bar{\psi}\bar{\varphi}^{-1}$ is $-\epsilon^{s-t-1}$-adjoint with respect to g.

Thus $\dim_{E_0} \ker D = \dim_{E_0} \mathbf{A}(-\epsilon^{s-t-1})$ in the terminology of Lemma 7. Suppose first that $\epsilon = -1$, i.e., $\bar{\sigma} = \pm 1$ and $E = E_0 = F$. Since g is $\epsilon^{s-1}\bar{\sigma}$-symmetric and the forms in $\mathbf{B}$ are $\epsilon^t\bar{\sigma}$-symmetric, we get by Lemma 7(b)

$$\dim \ker D + \dim \mathbf{B} = (1/2)\, n(n + \epsilon^{s-1}\bar{\sigma}(-\epsilon^{s-t-1})) + (1/2)\, n(n + \epsilon^t\bar{\sigma}) = n^2.$$

Now $\dim \ker D + \dim \operatorname{im} D = \dim \operatorname{Hom}(\bar{V}, \bar{W}) = n^2$, so $\operatorname{im} D = \mathbf{B}$, i.e. D is surjective. Now suppose $\epsilon = 1$. By Theorem 8, $\dim_{E_0} \mathbf{B} = \dim_{E_0} \mathbf{H}(\iota, \epsilon^t\bar{\sigma})$ and so by Lemma 7(a)

$$\dim_{E_0} \ker D + \dim_{E_0} \mathbf{B} = n^2 + n^2 = \dim_{E_0} \operatorname{Hom}_E(\bar{V}, \bar{W})$$

and again D is onto. Therefore D is surjective in all cases.

(5) Now define a bilinear form

$$D_0 : \bar{V} \times \bar{V} \to F, \qquad D_0(\bar{u}, \bar{v}) = B(\pi^t u, v) - C(\pi^t\varphi u, \varphi v).$$

It is well defined, for if $v + \pi w$ is another representative of $\bar{v}$, for example, then $B(\pi^t u, \pi w) - C(\pi^t\varphi u, \varphi\pi w) = B(\pi^{t+1}\epsilon u, w) - C(\pi^{t+1}\varphi\epsilon u, \varphi w) = 0$ by (10) which shows that the choice of v does not change $D_0(\bar{u}, \bar{v})$. An easy calculation shows $D_0 \in \mathbf{B}(\iota, \epsilon^t\bar{\sigma})$.

Therefore there is an $\psi \in \operatorname{Hom}_{F[\sigma]}(V, W)$ such that $D_0 = D_{\bar{\psi}}$. This means that

$$B(\pi^t u, v) = C(\pi^t\varphi u, \varphi v) + C(\pi^{s-1}\psi u, \varphi v) + \epsilon^{s-t-1}C(\pi^{s-1}\varphi u, \psi v).$$

Add in the term $C(\pi^{s-1}\psi u, \pi^{s-t-1}\psi v)$, which is 0 since $t < s - 1$ implies $\pi^{2(s-1)-t} = 0$, to get

$$B(\pi^t u, v) = C(\pi^t(\varphi u + \pi^{s-t-1}\psi u), \varphi v + \pi^{s-t-1}\psi v).$$

Thus replacement of φ by $\varphi + \pi^{s-t-1}\psi$ yields (9). Since $s - t - 1 > 0$, φ an isomorphism implies $\varphi + \pi^{s-t-1}\psi$ is an isomorphism (e.g., by determinants in $F[\sigma]$) and so Theorem 9 is proved.

4. Case IIb

Now we suppose that $p = X + 1$ and characteristic of $F = 2$. As prime element in $F[\sigma]$ we choose $\pi = \sigma + 1$. The adjoint of π is $\sigma^{-1} + 1 = \sigma^{-1}\pi$. So for example,

$$B(\pi u, v) = B(u, \sigma^{-1}\pi v).$$

As in Case IIa, we consider the bilinear form $B(\pi^{s-1}u, v)$ on $V(s) \times V(s)$

for $s = 1,..., r$ where $(X + 1)^r$ is the minimal polynomial of σ. Since $\pi^{s-1}\sigma = \pi^{s-1}$ on $V(s)$, it is easy to see that this form is symmetric. And it induces a non-degenerate symmetric bilinear form h_s on $\overline{V}_s \times \overline{V}_s$ where $\overline{V}_s$ is defined as before.

We next define quadratic forms

$$Q\pi^s: V \to F, \qquad s = 0, 1,...$$

by $Q\pi^s(v) = B(\pi^s v, \pi^s v)$ where $\pi^0 = 1$. The bilinear form $\beta\pi^s$ of $Q\pi^s$ is $\beta\pi^s(u, v) = B(\pi^s u, \pi^{s+1} v)$, and $Q\pi^s = 0$ if $2s \geqslant r$. The radical of $\beta\pi^s$ is $V(2s + 1)$ and the radical of $Q\pi^s$ is

$$\begin{aligned} V(2s + 1)_0 &= \{v \in V(2s + 1) : B(\pi^s v, \pi^s v) = 0\} \\ &= \{v \in V(2s + 1) : h_{2s+1}(\bar{v}, \bar{v}) = 0\}. \end{aligned}$$

Thus $\pi V(2s + 2) + V(2s) \subseteq V(2s + 1)_0$. Obviously $V(2s + 1)_0$ is an $F[\sigma]$-module.

To obtain a nondegenerate quadratic form, we divide out by $\operatorname{rad} Q\pi^s$: define Q_s to be the resulting quadratic form on $V/V(2s + 1)_0$, i.e., $Q_s(\bar{v}) = B(\pi^s v, \pi^s v)$. The defect of Q_s is $V(2s + 1)/V(2s + 1)_0$ so Q_s is nondefective iff h_{2s+1} is alternating.

Consider a second nondegenerate bilinear form $C: W \times W \to F$ with asymmetry similar to σ and denoted also by σ. Let $g_1, g_2,...$ be the symmetric forms attached to C and $R_0, R_1,...$ its quadratic forms.

THEOREM 10 (Case IIb). *$B \simeq C$ if and only if $h_{2s+1} \simeq g_{2s+1}$ and $Q_s \simeq R_s$ for $s = 0, 1,..., [r/2]$.*

Proof. The necessity is proved as in Theorem 9. Consider the sufficiency.

(1) We show first that it suffices to find a map $\varphi: V \to W$ such that

$$\varphi \text{ is an } F[\sigma]\text{-isomorphism and } R\pi^0(\varphi u) = Q\pi^0(u) \text{ for all } u \text{ in } V. \qquad (11)$$

Suppose that φ is such a mapping. It is also an isometry with respect to the forms $\beta\pi^0$ and $\gamma\pi^0$ (where $\gamma\pi^s$ is the bilinear form of $R\pi^s$) and hence

$$C(\varphi u, \varphi v) = B(u, v) \qquad (12)$$

if u or $v \in \pi V$.

(1a) We show that we may also suppose that φ induces an isometry $(V(1), B) \to (W(1), C)$, i.e., an isometry $V(1) \to W(1)$ with respect to $B \mid_{V(1)\times V(1)}$ and $C \mid_{W(1)\times W(1)}$. Take a splitting $V = V_1 \oplus V_2 \oplus \cdots \oplus V_r$ of the usual kind and let $V_0 = V_2 \oplus \cdots \oplus V_r$. Since φ is an $F[\sigma]$-iso-

morphism $W = \varphi V_1 \oplus \cdots \oplus \varphi V_r$ is a splitting of the same type and it is easy to check that the restriction of C to $\varphi V_1 = \varphi V_1/\pi\varphi V_1$ is equivalent to g_1 (cf. (7)). Similarly the restriction of B to V_1 is equivalent to h_1, and so there is an isometry $\varphi_1 : (V_1, B) \to (\varphi V_1, C)$. Since $\pi V_1 = 0 = \pi\varphi V_1$, it is trivially an $F[\sigma]$-homomorphism. Let $\varphi' = \varphi_1 \oplus \varphi \mid_{V_0}$. Since $\operatorname{im}\varphi' = \operatorname{im}\varphi = W$, φ' is an $F[\sigma]$-isomorphism. Now $V(1) = V_1 \oplus \pi V(2)$, see (4) e.g., and by (6) the radical of B on $V(1) \times V(1)$ is $\pi V(2)$; similar statements hold for W and it follows immediately that $\varphi' : (V(1), B) \to (W(1), C)$ is an isometry.

The equality $R\pi^0(\varphi' v) = Q\pi^0(v)$ holds if v is in V_1 (by the choice of φ_1) or in V_0 (since it holds for φ), and so it holds in $V = V_1 + V_0$ since $V_1 \subseteq \operatorname{rad}\beta\pi^0$ and $\varphi V_1 \subseteq \operatorname{rad}\gamma\pi^0$. Thus when we replace φ by φ', we get the desired result.

(1b) Next we show that we may suppose that (12) holds if either u or v is in $V(1) + \pi V$. It has already been shown to hold if u or v is in πV and (11) is satisfied; thus it suffices to adjust φ in such a way that (11), and (12) for $u \in V(1)$ and v arbitrary, both hold. We may in fact suppose that $u \in V_1$ since $V(1) + \pi V = V_1 + \pi V$.

The mapping $w \mapsto B(w, \ )$ is an isomorphism of V onto its dual V^* which carries $\pi V(2)$ onto the subspace of V^* which annihilates $(\pi V(2))^\perp = V(1) + \pi V$. If $u \in V_1$, the mapping $v \mapsto C(\varphi u, \varphi v) - B(u, v)$ is also in this subspace (by (1a) and the remark preceding (1a)). Thus there is a unique vector in $\pi V(2)$, call it ψu, such that

$$B(\psi u, v) = C(\varphi u, \varphi v) - B(u, v) \tag{13}$$

for all v in V. It is easy to see that $\psi : V_1 \to \pi V(2)$ is an F-homomorphism, and hence also an $F[\sigma]$-homomorphism. We may suppose that ψ is actually an $F[\sigma]$-homomorphism of V into $\pi V(2)$ by defining it to be 0 on $V_2 \oplus \cdots \oplus V_r$. The $F[\sigma]$-homomorphism $1 + \psi$ is actually an isomorphism since if $(1 + \psi)u = 0$ then $u = \psi u$ is in $\pi V(2) \subseteq \ker\psi$ so $u = \psi u = 0$.

We shall now show that $\varphi' = \varphi(1 + \psi)$ has the required properties, enabling us to replace φ by φ'.

Since $\operatorname{im}\varphi\psi \subseteq \pi W(2) \subseteq \operatorname{rad} R\pi^0$, $R\pi^0(\varphi' v) = R\pi^0(\varphi v) = Q\pi^0(v)$. Similarly since $\operatorname{im}\varphi\psi \subseteq W(1)^\perp$, $C(\varphi' u, \varphi' v) = C(\varphi' u, \varphi v)$ if $u \in V(1)$; since $\operatorname{im}\psi \subseteq \pi V$, $C(\varphi\psi u, \varphi v) = B(\psi u, v)$ and so $C(\varphi' u, \varphi v) = B(u, v)$ if $u \in V_1$ when one takes (13) into account.

(1c) Conclusion of proof of (1). The subspace Ψ of $\operatorname{Hom}_F(W, \pi W(2))$ of mappings whose kernel contains $W' = W(1) + \pi W$ is isomorphic to $\operatorname{Hom}_F(W/W', \pi W(2))$ and consists of $F[\sigma]$-homomorphisms since the kernels contain πW and the images are annihilated by π. Since $(\pi W(2))^\perp = W'$,

$\dim_F W/W' = \dim_F \pi W(2) = m$ say, so $\dim_F \Psi = m^2$. If $\psi \in \Psi$, the bilinear form $C(\psi u, v)$ on W has left radical $\ker \psi \supseteq W'$ and right radical $= (\operatorname{im} \psi)^\perp$ which contains $(\pi W(2))^\perp = W'$; we therefore obtain a bilinear form C_ψ on W/W' given by $C_\psi(\bar{u}, \bar{v}) = C(\psi u, v)$. The mapping $\psi \mapsto C_\psi$ is clearly F-linear and injective and hence by dimensions must be onto the space of bilinear forms of W/W'. Choose ψ so that $D = C_\psi$ is symmetric and non-degenerate.

If $\bar{\theta} \in \operatorname{End}_F W/W'$, let $\Delta(\bar{\theta})$ be the bilinear form $D(\bar{\theta}\bar{u}, \bar{v}) + D(\bar{u}, \bar{\theta}\bar{v})$ on W/W'. Then Δ is an F-linear mapping of $\operatorname{End} W/W'$ into the space of alternating forms on W/W' which has dimension $(1/2)\, m(m-1)$. The kernel of Δ consists of those $\bar{\theta}$ which are self adjoint with respect to D and therefore has dimension $(1/2)\, m(m+1)$ by Lemma 7. Δ is therefore surjective, i.e., every alternating form on W/W' is of the form $\Delta(\bar{\theta})$.

Since φ is an isometry with respect to $Q\pi^0$ and $R\pi^0$, the bilinear form $C(u, v) - B(\varphi^{-1}u, \varphi^{-1}v)$ is alternating; since its radical contains W' by (1b), it induces an alternating form on W/W' which is of the form $\Delta(\bar{\theta})$ for some $\bar{\theta}$. Let $\theta \in \operatorname{End}_{F[\sigma]} W$ be a lifting of this $\bar{\theta}$. Since $D(\bar{u}, \bar{\theta}\bar{v}) = D(\bar{\theta}\bar{v}, \bar{u})$, we get

$$C(u, v) - B(\varphi^{-1}u, \varphi^{-1}v) = C(\psi\theta u, v) + C(\psi\theta v, u).$$

Since $\operatorname{im} \psi \subseteq W(1)$, $\sigma\psi = \psi$ and so $C(\psi\theta v, u) = C(u, \psi\theta v)$; thus upon replacing u and v by φu and φv and setting $(1 + \psi\theta)\varphi = \varphi'$, the above equation becomes

$$B(u, v) = C(\varphi' u, \varphi' v)$$

for all u and v in V since $\operatorname{im} \psi \subseteq \pi W(2)$ implies $C(\psi\theta\varphi u, \psi\theta\varphi v) = 0$. This finishes the proof of (1).

(2) We now prove the following lemma: let M and N be isomorphic finitely generated $F[\sigma]$-modules; then any $F[\sigma]$-isomorphism $\psi\colon M/M(s) \to N/N(s)$ can be "lifted" to an $F[\sigma]$-isomorphism $\Psi\colon M \to N$.

Proof. Write M as a direct sum of cyclic modules and let M_1 (resp. M_2) be the direct sum of those cyclic modules which are (resp. are not) annihilated by π^s. Then $M = M_1 \oplus M_2$, M_1 (resp. M_2) maps onto 0 (resp. $M/M(s)$) under the canonical map $M \to M/M(s)$, and $M_2(s) \subseteq \pi M_2$. Write $N = N_1 \oplus N_2$ in the same manner. Then $M_1 \simeq N_1$, $M_2 \simeq N_2$; let $\Psi_1 : M_1 \simeq N_1$ be any $F[\sigma]$-isomorphism.

Suppose that M_2 is the direct sum of the nonzero cyclic modules $F[\sigma]\, x_1, \ldots, F[\sigma]\, x_q$ and choose $y_1, \ldots, y_q$ in N_2 such that the image of y_i in $N/N(s)$ is the image of x_i under $M \to M/M(s) \xrightarrow{\psi} N/N(s)$ for $i = 1, \ldots, q$. Let $\Psi_2 : M_2 \to N_2$ be the unique $F[\sigma]$-homomorphism such that $\Psi_2 x_i = y_i$

for $i = 1,..., q$. Then $\Psi = \Psi_1 \oplus \Psi_2 : M \to N$ is clearly a lifting of ψ and to show it is an isomorphism it suffices to show that Ψ_2 is an isomorphism. Since the x_i generate M_2, their images in $N/N(s)$ generate it, so their images in $(N/N(s))/\pi(N/N(s))$ generate it. But this latter module is isomorphic to $(N_2/N_2(s))/\pi(N_2/N_2(s))$ which is isomorphic to $N_2/\pi N_2$ since $N_2(s) \subseteq \pi N_2$. These isomorphisms are all "natural" and one infers without difficulty that the images of the y_i under $N_2 \to N_2/\pi N_2$ generate this latter module, and so (e.g., by Nakayama's lemma [2]) the y_i generate N_2, which implies that Ψ_2 is onto, and hence an isomorphism (say by Jordan-Hölder lengths).

(3) The theorem can now be proved by induction on r. When $r = 1$, $B = h_1 \simeq g_1 = C$. Suppose then that $r > 1$. The bilinear form $B(\pi u, \pi v)$ on V has radical $V(2)$ and so defines a nondegenerate bilinear form B' on $V' = V/V(2)$, $B'(u', v') = B(\pi u, \pi v)$. Its asymmetry σ' and $\pi' = 1 + \sigma'$ are the transformations induced on V' by σ and π. Since

$$V'(s) = V(s + 2)/V(2) \qquad \text{for} \quad s \geqslant 0,$$

there is a natural isomorphism between $V'(s)/\pi' V'(s + 1) + V'(s - 1)$ and $V(s + 2)/\pi V(s + 3) + V(s + 1)$ for $s \geqslant 1$, given by $\bar{v}' \mapsto \bar{v}$ where $v \in V(s + 2)$. It is actually an isometry: $h_s'(\bar{u}', \bar{v}') = B'(\pi'^{s-1}u', v') = B(\pi^s u, \pi v)$; since u and $v \in V(s + 2)$, this is equal to $B(\pi^{s+1}u, v) = h_{s+2}(\bar{u}, \bar{v})$.

We next show that Q_s' and Q_{s+1} are also equivalent. First of all $V'(2s + 1)_0$ consists of the v' in $V'(2s + 1)$ such that $h'_{2s+1}(\bar{v}', \bar{v}') = 0$, so by the above it is the image in V' of those $v \in V(2s + 3)$ with $h_{2s+3}(\bar{v}, \bar{v}) = 0$, i.e., $V'(2s + 1)_0 = V(2s + 3)_0/V(2)$. Therefore there is a natural isomorphism between $V'/V'(2s + 1)_0$ and $V/V(2s + 3)_0$ and it is easily checked that it is an isometry with respect to Q_s' and Q_{s+1}.

In a similar manner we may define a bilinear form C' on $W' = W/W(2)$ with associated forms $g_1',..., R_1',...$. Then $g'_{2s+1} \simeq g_{2s+3} \simeq h_{2s+3} \simeq h'_{2s+1}$ and similarly $R_s' \simeq Q_s'$. Since V' and W' are isomorphic as $F[\sigma]$-modules they are also isomorphic as $F[\sigma']$-modules so the asymmetries of V' and W' are similar. Therefore by the induction hypothesis there is an isometry θ': $V' \to W'$, $C'(\theta' u', \theta' v') = B'(u', v')$ which is necessarily $F[\sigma']$-linear (cf. proof of Theorem 4) and so also $F[\sigma]$-linear.

By (2), θ' can be lifted to an $F[\sigma]$-isomorphism θ: $V \to W$. If $V = V_1 \oplus \cdots \oplus V_r$ is a splitting of the usual kind, the restrictions of B and C to V_1 and $W_1 = \theta V_1$ resp. are equivalent to h_1 and g_1 resp., and so we may change $\theta \mid_{V_1}$ to ensure that it is an isometry on V_1 with respect to B and C; since $V_1 \subseteq V(2)$ and $W_1 \subseteq W(2)$, the new θ will still lift θ', and will still be $F[\sigma]$-linear since σ is 1 on V_1 and W_1.

It follows that θ is also an isometry on each of the subspaces $V(1)$ and πV, with respect to B and C. Indeed $V(1) = V_1 \oplus \pi V(2)$ and $W(1) =$

$W_1 \oplus \pi W(2)$, and $\pi V(2)$ and $\pi W(2)$ are the radicals of B and C restricted to $V(1)$ and $W(1)$, resp. As for πV, one uses the facts that θ' is an isometry and θ is $F[\sigma]$-linear.

Now $\theta V(1) = W(1)$, and if $v \in V(1)$ then $B(v, v) = 0$ iff $C(\theta v, \theta v) = 0$, so $\theta V(1)_0 = W(1)_0$ and θ induces an $F[\sigma]$-isomorphism

$$\theta^*\colon V^* = V/V(1)_0 \to W^* = W/W(1)_0 .$$

It satisfies $R_0(\theta^*\pi v^*) = C(\theta\pi v, \theta\pi v) = B(\pi v, \pi v) = Q_0(\pi v^*)$ and so by Witt's theorem (see the appendix) $\theta^*|_{\pi V^*}$ can be extended to an isometry $\varphi^*\colon (V^*, Q_0) \to (W^*, R_0)$ since πV^* and πW^* have intersection 0 with the defects of Q_0 and R_0 (the defect of Q_0, for example, is the image of V_1 in V^*).

Let β_0 and γ_0 be the bilinear forms belonging to Q_0 and R_0. An easy calculation, using $\gamma_0(u^*, v^*) = C(u, \pi v)$ and the fact that θ is an $F[\sigma]$-isometry on πV, shows that $\gamma_0(\theta^*\pi u^*, \theta^* v^*) = \beta_0(\pi u^*, v^*)$, and from this one gets $\gamma_0(\theta^*\pi u^*, (\varphi^* - \theta^*) v^*) = 0$ for all u and v in V. This means that $\pi(\varphi^* - \theta^*) V^* \subseteq \operatorname{def} R_0 \cap \pi W^* = 0$ i.e., $\pi(\varphi^* - \theta^*) = 0$. Since

$$(\varphi^* - \theta^*)\pi = 0 \qquad \text{and} \qquad \theta^*\pi = \pi\theta^*,$$

we get $\pi\varphi^* = \varphi^*\pi$, i.e., φ^* is $F[\sigma]$-linear.

We show next that φ^* can be lifted to an $F[\sigma]$-isomorphism $V \to W$. Choose an $F[\sigma]$-splitting $V^* = \operatorname{def} Q_0 \oplus V_0^*$, and then a subspace V' of V which maps isomorphically onto $\operatorname{def} Q_0$ under the canonical map $V \to V^*$; the map $V' \to \operatorname{def} Q_0$ is an $F[\sigma]$-isomorphism since σ is the identity on both (the inverse image of $\operatorname{def} Q_0$ in V is $V(1)$). If V_0 is the inverse image of V_0^* in V, it follows that $V = V' \oplus V_0$ and $V_0/V_0(1) \simeq V_0^*$ since $V_0(1) = V(1)_0$. Now $\varphi^* \operatorname{def} Q_0 = \operatorname{def} R_0$ so $W^* = \operatorname{def} R_0 \oplus W_0^*$ where $W_0^* = \varphi^* V_0^*$. The analogous construction for W leads to $W = W' \oplus W_0$. Clearly V' and W' are $F[\sigma]$-isomorphic and therefore so are V_0 and W_0 by elementary divisor theory. The required lifting φ of φ^* is obtained as the "direct sum" of the (unique) lifting of $\varphi^*|_{\operatorname{def} Q_0}$ to $V' \to W'$, and a lifting of $\varphi^*|_{V_0^*}$ to $V_0 \to W_0$ which exists by (2).

It is clear that φ is an isometry with respect to $Q\pi^0$ and $R\pi^0$, and so by (1) the theorem is completely proved.

THEOREM 11. *Let F be an algebraically closed field of arbitrary characteristic. Then two nondegenerate bilinear forms over F are equivalent if and only if their asymmetries are similar and, in the case of char $F = 2$ and $p = X + 1$ being a divisor of the minimal polynomial of the asymmetries, the symmetric forms h_{2s+1} and g_{2s+1} arising from p are both alternating or both nonalternating for $s = 0, 1, \ldots$.*

Proof. The necessity is part of Theorems 4 and 9. Conversely, if both forms are in Case I, Theorem 5 applies. Suppose they are both in Case II, so $p = X \pm 1$. Since the asymmetries are similar, V and W are isomorphic as modules, so $\overline{V}_s$ and $\overline{W}_s$ have the same dimension over F for each s. Thus h_s and g_s have the same rank. In case IIa, i.e., $\operatorname{char} F \neq 2$, they are both $(-1)^{s-1}\bar{\sigma}$-symmetric, i.e., they are both alternating or both symmetric, and so are equivalent; therefore B and C are equivalent by Theorem 9. In Case IIb, $\operatorname{char} F = 2$. Since a nonalternating symmetric form can be diagonalized [1], any two (of the same rank) are equivalent, as is the case for alternating forms. The defect of Q_s has dimension 0 or 1 according to whether h_{2s+1} is alternating or not (see preamble of Theorem 10), and Q_s is a sum of hyperbolic forms and its defect; since the number of variables of the nondefective part of Q_s is the dimension of $V/V(2s+1)$, it follows easily that Q_s and R_s are equivalent, so the bilinear forms are equivalent by Theorem 10. The general case of Theorem 11 now follows from Theorem 4.

Remark. When applying Theorem 11 in characteristic 2, one can use the fact that a nondegenerate alternating form has even rank, and so a nondegenerate symmetric form of odd rank cannot be alternating.

Appendix

Let D be a division ring. Suppose that ι is an antiautomorphism of D and μ an element of D which satisfy

$$\mu\mu^{\iota} = 1, \qquad a^{\iota^2} = \mu a \mu^{-1} \qquad \text{for all} \quad a \text{ in } D.$$

Define $D_{\iota,\mu} = \{a - a^{\iota}\mu : a \in D\}$ and $D^{(\iota,\mu)} = D/D_{\iota,\mu}$ (additive factor group). D acts on $D^{(\iota,\mu)}$ via $b \,.\, \bar{a} = \overline{a^{\iota}ba}$. We shall assume that $\mu \neq -1$ if ι is the identity and $\operatorname{char} D \neq 2$.

Let V be a right vector space (of finite dimension) over D. An (ι, μ)-*quadratic form*, or a *pseudoquadratic form with respect to* ι *and* μ, is a map $Q: V \to D^{(\iota,\mu)}$ which satisfies

(i) $Q(va) = Q(v) \,.\, a$ for all v in V and a in D,

(ii) $Q(u+v) = Q(u) + Q(v) + \overline{f(u,v)}$ for all u and v in V,

for some trace valued (ι, μ)-hermitian form $f: V \times V \to D$.

The conditions of f mean that f is biadditive, $f(ua, vb) = a^{\iota} f(u,v) b$, $f(v,u) = f(u,v)^{\iota}\mu$, and that for each u in V, $f(u,u) = a + a^{\iota}\mu$ for some a in D. An account of the basic properties of pseudoquadratic forms can be found in [8]. Note that Q is a quadratic form if ι is the identity and $\mu = 1$.

The form Q is called nondegenerate if $Q(u + v) = Q(u)$ for all u implies that $v = 0$. This means that Q is an injective homomorphism on the defect $V^\perp = \{v \in V : f(v, V) = 0\}$.

The proof of the following version of Witt's theorem is an adaptation of a proof of G. E. Wall [9] for the case of hermitian forms; we shall show in a corollary how one can deduce Witt's theorem for trace-valued hermitian forms (other than alternating forms in characteristic $\neq 2$) from it in a formal manner.

THEOREM 12. *Let φ: $U \to V$ be a D-monomorphism of the subspace U of V satisfying $Q(\varphi u) = Q(u)$ for all u in U. Then if Q is nondegenerate, φ can be extended to an element of the orthogonal group of Q if and only if $\varphi(U \cap V^\perp) = (\varphi U) \cap V^\perp$.*

Proof. The necessity follows from the fact that any element of the orthogonal group $O(Q)$ is the identity on $V^\perp$. Similarly, $\varphi(U \cap V^\perp) \subseteq V^\perp$ and $Q(\varphi u) = Q(u)$ imply that $\varphi\,|_{U \cap V^\perp} = 1$. To prove the sufficiency we may suppose that $V^\perp \subseteq U$ since the map $v + u \mapsto v + \varphi u$, where $v \in V^\perp$ and $u \in U$, yields a well defined extension of φ to $V^\perp + U$. The proof proceeds by induction on $m = \dim U \geqslant \dim V^\perp$.

If $m = \dim V^\perp$, φ is the identity on U and hence is extended by the identity on V. So suppose that $m > \dim V^\perp$. Let H be a hyperplane of U containing $V^\perp$. The induction hypothesis applies to $\varphi\,|_H$ so we can find $\psi \in O(Q)$ extending it. Any extension of $\psi^{-1}\varphi$ when composed with ψ will be an extension of φ; therefore we may suppose that

$$\varphi\,|_H = 1.$$

Suppose that $U = H \oplus u_1 D$ and let $u_0 = u_1 - \varphi u_1$. Then

$$u_0 \in H^\perp,\ Q(u_0) = \overline{f(u_0, u_1)}, \qquad f(u_0, u_0) = f(u_0, u_1) + f(u_1, u_0). \quad (14)$$

The second equation follows from $Q(u_1) = Q(\varphi u_1) = Q(-u_0 + u_1) = Q(u_0) + Q(u_1) - \overline{f(u_0, u_1)}$, and the first and third are also easy; note that $f(\varphi u, \varphi v) = f(u, v)$ for u and v in U since the two sides of the equation are the uniquely determined hermitian forms of $Q(\varphi u)$ and $Q(u)$.

Suppose first that $f(u_0, u_1) \neq 0$. Define

$$\hat{\varphi} v = v - u_0 f(u_0, u_1)^{-1} f(u_0, v).$$

$\hat{\varphi}$: $V \to V$ is clearly D-linear and $Q(\hat{\varphi} v)$ is equal to $Q(v)$ plus the terms $Q(u_0) \cdot (f(u_0, u_1)^{-1} f(u_0, v))$ and the image of $-f(u_0, v)^{\iota} f(u_0, u_1)^{-\iota} f(u_0, v)$ in $D^{(\iota,\mu)}$. But the first of these terms is the image in $D^{(\iota,\mu)}$ of

$$f(u_0, v)^{\iota} f(u_0, u_1)^{-\iota} f(u_0, u_1) f(u_0, u_1)^{-1} f(u_0, v)$$

by (14) and so $Q(\hat{\varphi}v) = Q(v)$, i.e., $\hat{\varphi} \in O(Q)$. Furthermore, $\hat{\varphi}u = u$ if $u \in H$ and $\hat{\varphi}u_1 = u_1 - u_0 = \varphi u_1$ so $\hat{\varphi}$ is the required extension of φ.

Suppose now that $f(u_0, u_1) = 0$. Then

$$Q(u_0) = 0, \qquad f(u_0, u_0) = 0$$

by (14). We next show that there is a vector u_2 in V satisfying

$$u_2 \in H^{\perp}, \qquad f(u_2, u_1) = 1 \neq f(u_2, u_0). \tag{15}$$

If $u_1 \notin H + u_0 D$ then $(H + u_0 D)^{\perp}$ is not contained in $(u_1 D)^{\perp}$ so u_2 can be chosen to satisfy (15) and $f(u_2, u_0) = 0$. If $u_1 \in H + u_0 D$ then $u_0 = u + u_1 a$ with $u \in H$ and $a \in D$. Since $\varphi u_1 \notin H$, $a \neq 1$, so any u_2 in $H^{\perp}$ with $f(u_2, u_1) = 1$ satisfies (15) since $f(u_2, u_0) = a$.

Now $f(u_0, u_2) = f(u_2, u_0)^{\iota}\mu \neq \mu$ so we may choose b and c in D to satisfy

$$1 + f(u_0, u_2)b = \mu b,$$

$$\text{image in } D^{(\iota,\mu)} \text{ of } c^{\iota}(1 + f(u_0, u_2)b) = Q(u_2 b).$$

A direct verification will show that

$$\hat{\varphi}v = v + u_2 b f(u_0, v) - u_0 f(u_2, v) - u_0 c f(u_0, v) \tag{16}$$

is the required extension of φ. Let v, t_1, t_2 and t_3 denote the four terms, in order, of the right side. In calculating $Q(\hat{\varphi}v)$, one gets $Q(v) + Q(u_2 b) \cdot f(u_0, v)$ plus the images in $D^{(\iota,\mu)}$ of $f(v, t_1)$, $f(t_2, v)$, $f(t_3, v)$, $f(t_2, t_1)$ and $f(t_3, t_1)$ (all other possible terms are 0). When one writes

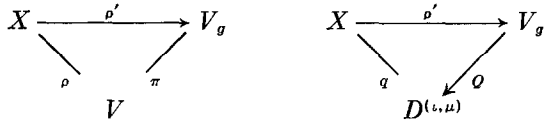
$$f(v, t_1) = f(v, u_2)\, b f(u_0, v) = f(u_2, v)^{\iota}\, \mu b f(u_0, v),$$

the equality $Q(\hat{\varphi}v) = Q(v)$ follows without difficulty.

COROLLARY. *Let $g: V \times V \to D$ be a nondegenerate trace-valued (ι, μ)-hermitian form and assume g is not alternating if* char $D \neq 2$. *Then any D-monomorphism $\varphi: U \to V$ such that $g(\varphi u, \varphi v) = g(u, v)$ for all u and v in U has an extension in the unitary group of g.*

Proof. Consider triples (X, q, ρ) consisting of a vector space X over D, an (ι, μ)-quadratic form $q: X \to D^{(\iota,\mu)}$, and a D-homomorphism $\rho: X \to V$ such that $g(\rho x, \rho y)$ is the hermitian form of q. In [8] Tits shows the existence of a "universal" such triple (V_g, Q, π); it is uniquely determined by the

following universal mapping property: if (X, q, ρ) is a triple as above, there is a unique D-homomorphism $\rho'\colon X \to V_g$ making both of



commutative. The existence of the universal triple does not depend on the nondegeneracy of g.

Since g is trace-valued V itself has a pseudoquadratic form with hermitian form g and it follows that π is onto, and then that $\ker \pi$ is the defect $V_g^{\perp}$ of Q. Moreover Q is nondegenerate since if $u \in \ker \pi$ and $Q(u) = 0$, the uniqueness requirement of the universal mapping property implies that the inclusion $uD \to V_g$ is the 0 map.

If W is any subspace of V (endowed with the restriction of g) it is evident that $\pi^{-1}W$, along with the restrictions of π and Q, is the universal object W_g for W. Thus if $W = \varphi U$, the map $\varphi \circ (\pi|_{U_g})\colon U_g \to W$ lifts to a map $\varphi'\colon U_g \to W_g$ preserving Q. There is a similar map $(\varphi^{-1})'\colon W_g \to U_g$ which by the uniqueness property must be the inverse of φ' so φ' is an isomorphism, and $\varphi' \ker \pi = \ker \pi$. We may therefore apply the theorem and obtain an extension $\Phi' \in O(Q)$ of φ'. Since Φ' is also in the unitary group of the hermitian form belonging to Q, the map Φ it induces on V is in the unitary group of g and is the desired extension of φ.

References

1. A. A. Albert, Symmetric and alternate matrices in an arbitrary field I, *Trans. Amer. Math. Math. Soc.* **43** (1938), 386–436.
2. N. Bourbaki, "Algèbre," Chapter VIII, Act. Sci. et Ind. 1261, Hermann, Paris, 1958.
3. I. K. Cikunov, The structure of isometric transformations of a symplectic or orthogonal vector space, *Sov. Math. Dokl.* **6** (1965), 1479–1481.
4. A. Fröhlich and A. M. McEvett, Forms over rings with involution, *J. Algebra* **12** (1969), 79–104.
5. J. Milnor, On isometries of inner product spaces, *Invent. Math.* **8** (1969), 83–97.
6. O. T. O'Meara, Integral equivalence of quadratic forms in ramified local fields, *Amer. J. Math.* **79** (1957), 157–186.
7. T. Springer, "Over Symplectische Transformaties," Thesis, Univ. of Leiden, 1951.
8. J. Tits, Buildings of spherical type and finite *BN*-pairs, *in* "Lecture Notes in Mathematics," Springer-Verlag, Berlin/Heidelberg/New York, to appear.
9. G. E. Wall, On the conjugacy classes in the unitary, symplectic and orthogonal groups, *J. Austral. Math. Soc.* **III** (1963), 1–62.

10. J. WILLIAMSON, On the algebraic problem concerning the normal forms of linear dynamical systems, *Amer. J. Math.* **58** (1936), 141–163.
11. H. J. ZASSENHAUS, On a normal form of the orthogonal transformation, *Canad. Math. Bull.* 1 (1958), 31–39, 101–111, 183–191.